

Week 3

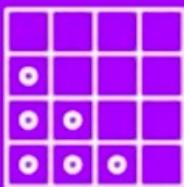
# Waffle Charts and Word Cloud

---

© IBM Corporation. All rights reserved.

# What you will learn

---



Explore waffle charts with the help of an illustration



Identify the use cases of waffle charts



Explore word cloud with the help of an illustration



Identify the use cases of word cloud

# Waffle charts

---

- Represents categorical data in the form of:
  - Square tiles
  - Cells
- Displays proportion or percentage of different categories
- Simplifies data for all types of audiences

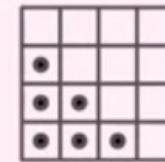


# Waffle charts: Use case

---

You can use waffle charts for:

- Market share analysis
- Demographic representation
- Project progress tracking
- Budget allocation
- Survey responses
- Election results
- Product sales analysis



# Waffle charts: Use case

## Market share analysis

- Visualize market share data

## Demographic representation

- Display demographic data

## Project progress tracking

- Represent completion status of tasks or milestones

## Budget allocation

- Demonstrate allocation of budgetary resources

# Waffle charts: Use case

## Survey responses

- Summarize survey responses

## Election results

- Provide a clear visualization of voting outcomes

## Product sales analysis

- Illustrate product sales

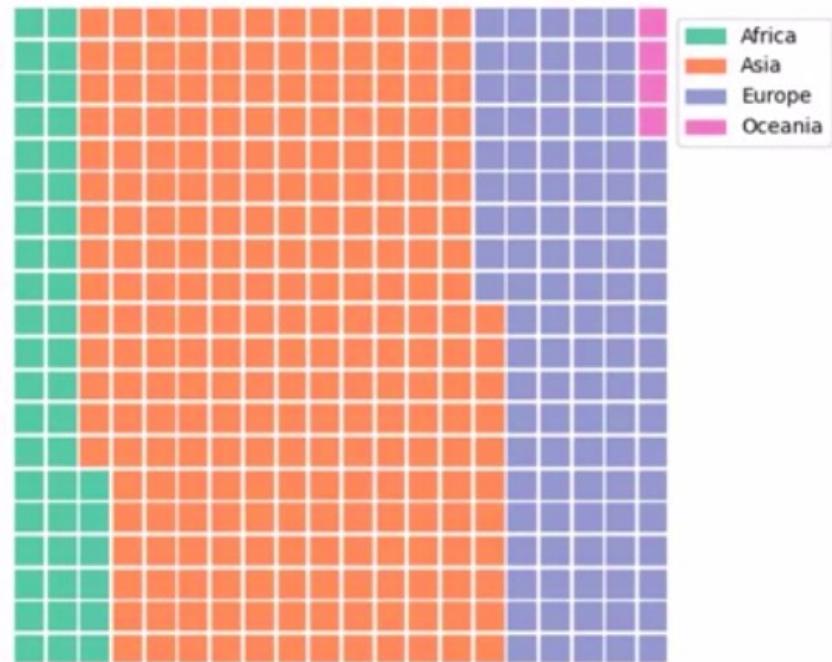
# pywaffle library

```
import matplotlib.pyplot as plt
from pywaffle import Waffle

# Create data for the waffle chart
data = df_dsn[['Africa', 'Asia', 'Europe', 'Oceania']]

# Set up the waffle chart figure
fig = plt.figure(
    FigureClass=Waffle,
    rows=20,
    columns=20,
    values=data,
    legend={'labels': ['Africa', 'Asia', 'Europe', 'Oceania'],
            'loc': 'upper left', 'bbox_to_anchor': (1, 1)})

# Display the chart
plt.show()
```



# Word cloud

---

- Is a popular data visualization method
- Presents a concise summary of textual content
- Depicts the importance of different words

# Word cloud

Bandusky 2

After leaving Reavis, I must attend college because it is definitely a requirement for becoming a veterinarian. In fact, a bachelor's degree is necessary in order to even enter a veterinarian program. One must also possess excellent communication, leadership, public speaking, and organizational skills. I have put a lot of thought and consideration into college, and I have decided that I would like to go to the University of Illinois. It is a wonderful school, and they even have a graduate program designed for students who want to become veterinarians.

Once I have completed a veterinarian program, I will be able to pursue my dream career. This career provides numerous benefits, the first of which is salary. The average veterinarian salary is \$60,000 a year, a salary that would definitely allow me to live a comfortable life. Secondly, it is a rewarding job. This job would provide me with the satisfaction of knowing that I am helping or saving an animal's life. Finally, becoming a veterinarian would assure me a lifetime of happiness. I know I would love going to my job every day, because I would be working with what I love most: animals.

In summary, when I graduate from Reavis, I plan to go to college to become a veterinarian. I love animals and I want to do anything that I can to help them. I know I am only a freshman, but I also know that I am growing up quickly. As Ferris Bueller quotes, "Life moves pretty fast. If you don't stop and look around once in a while, you could miss it!"



# Word cloud: Use case

---

You can use word cloud for:

- Social media analysis
- Customer feedback analysis
- Content analysis
- Market research
- Resume or job description analysis



# Word cloud: Use case

---



## Social media analysis

- Extract and visualize popular topics or sentiments



## Customer feedback analysis

- Summarize customer reviews or feedback



## Content analysis

- Analyze textual content

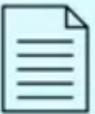
# Word cloud: Use case

---



## Market research

- Analyze survey responses



## Resume or job description analysis

- Highlight important skills or keywords in resumes

# Recap

---

In this video, you learned that:

- Waffle charts are a visualization technique that represent categorical data in the form of square tiles or cells.
- There are different areas in which you can use waffle charts.
- Word cloud is a popular data visualization method to visually present textual data in an engaging and informative manner.
- You can use word cloud in different areas to visually present textual data in an engaging and informative manner.

# Seaborn and Regression Plots

---

© IBM Corporation. All rights reserved.

# What you will learn



Explain what  
Seaborn is and what  
it does

# What you will learn

---



Explain what Seaborn is and what it does



Describe the functions of Seaborn

# Seaborn

---

- Seaborn is a Python visualization library based on Matplotlib.
- Seaborn offers built-in themes and color palettes to improve your plot visuals.
- Visuals that need ~20 lines of code using Matplotlib to be created, with Seaborn, the number of lines of code is reduced by 5-fold.

# Seaborn – Special Plots

---

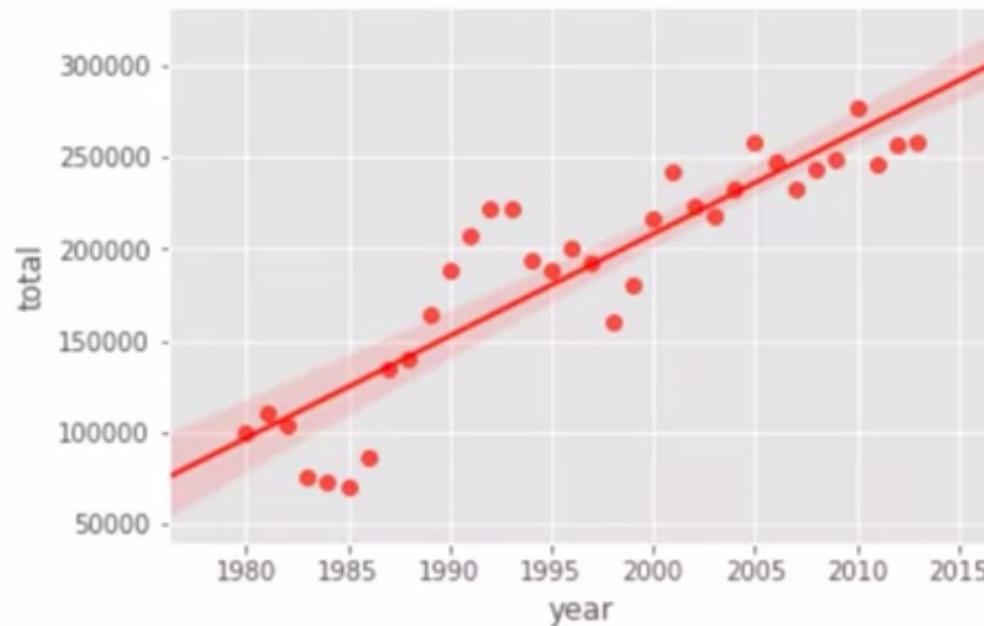
- Regression plot
- Distribution plot
- Categorical plot
- Creating Scatter plots with Regression lines.

# Regression Plots

df\_total

Year	Total
1980	99137
1981	110563
1982	104271
1983	75550
1984	73417
.	.
.	.

```
import seaborn as sns  
ax = sns.regplot(x='year', y='total', data=df_tot)
```

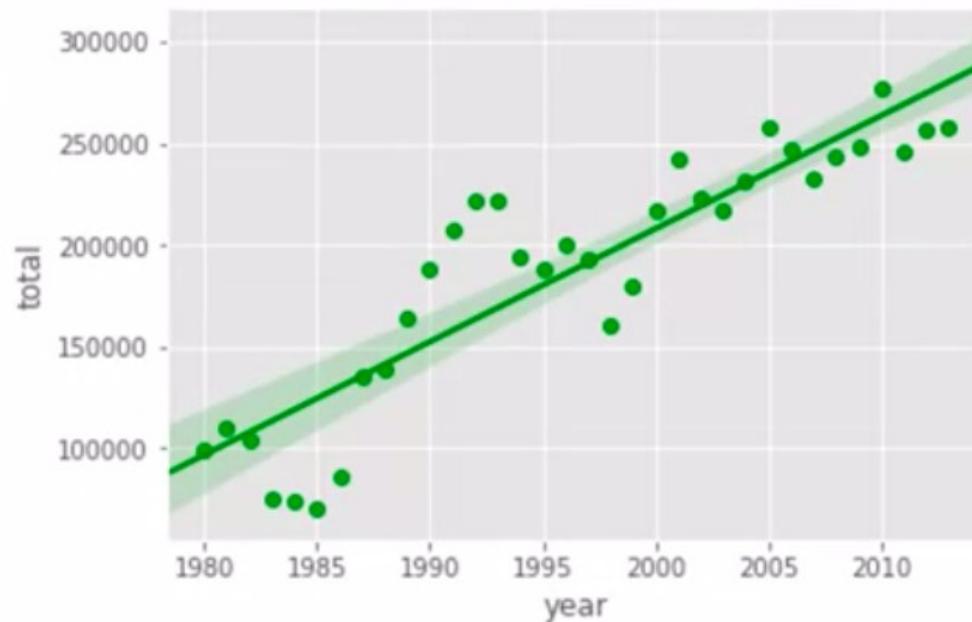


# Regression Plots

df\_total

Year	Total
1980	99137
1981	110563
1982	104271
1983	75550
1984	73417
.	.
.	.

```
import seaborn as sns  
ax = sns.regplot(x='year', y='total', data=df_tot,  
                  color='green')
```

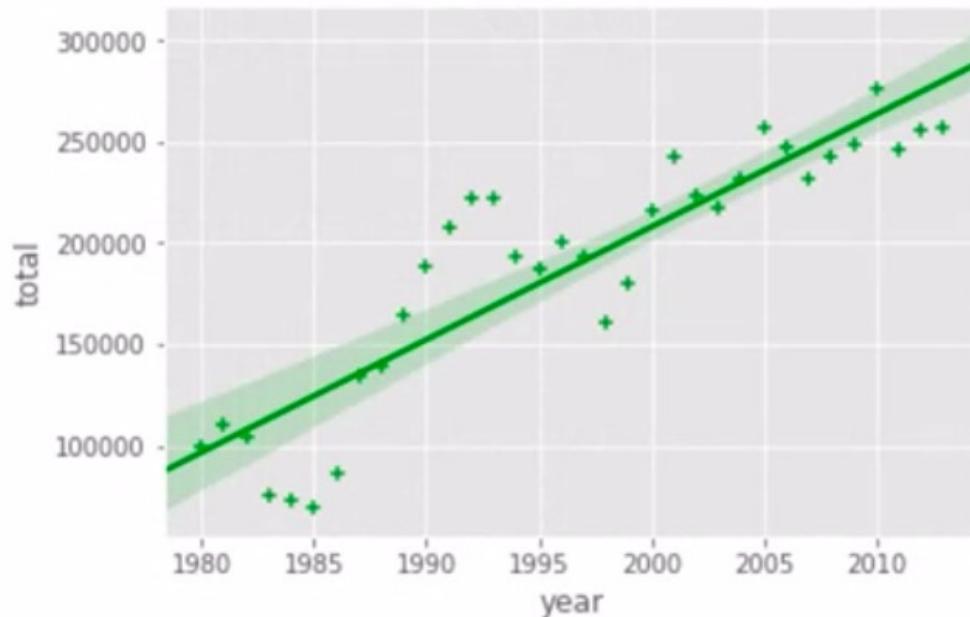


# Regression Plots

df\_total

Year	Total
1980	99137
1981	110563
1982	104271
1983	75550
1984	73417
.	.
.	.

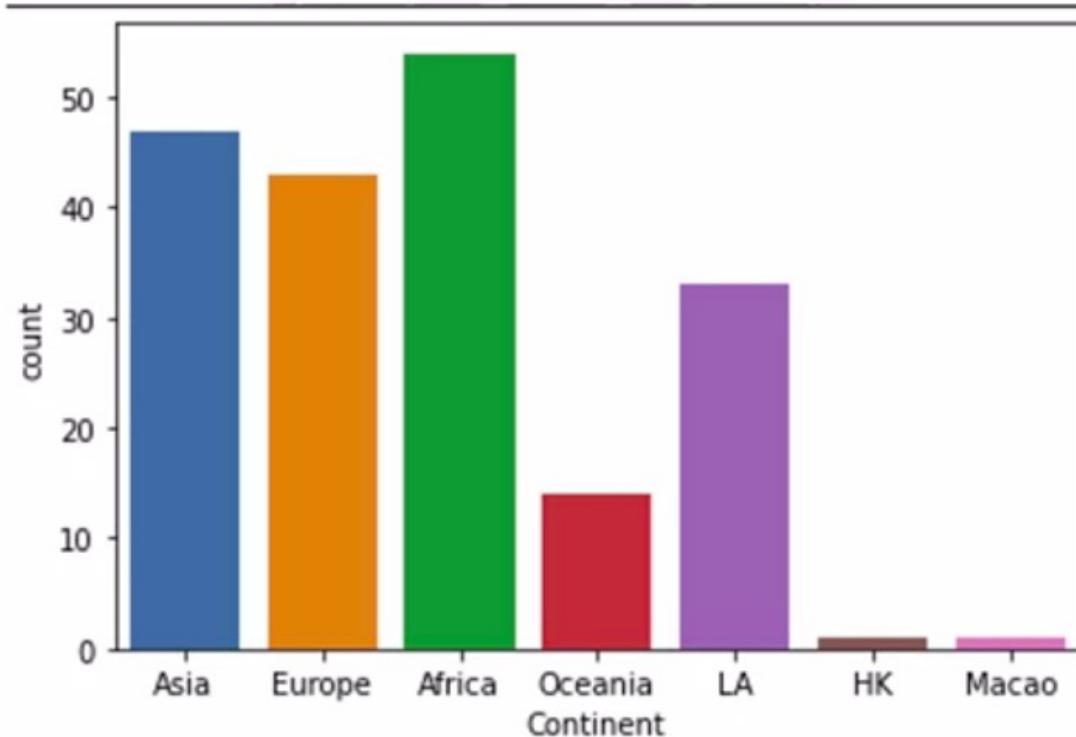
```
import seaborn as sns  
ax = sns.regplot(x='year', y='total', data=df_tot,  
                  color='green', marker='+')
```



# Categorical Plots

---

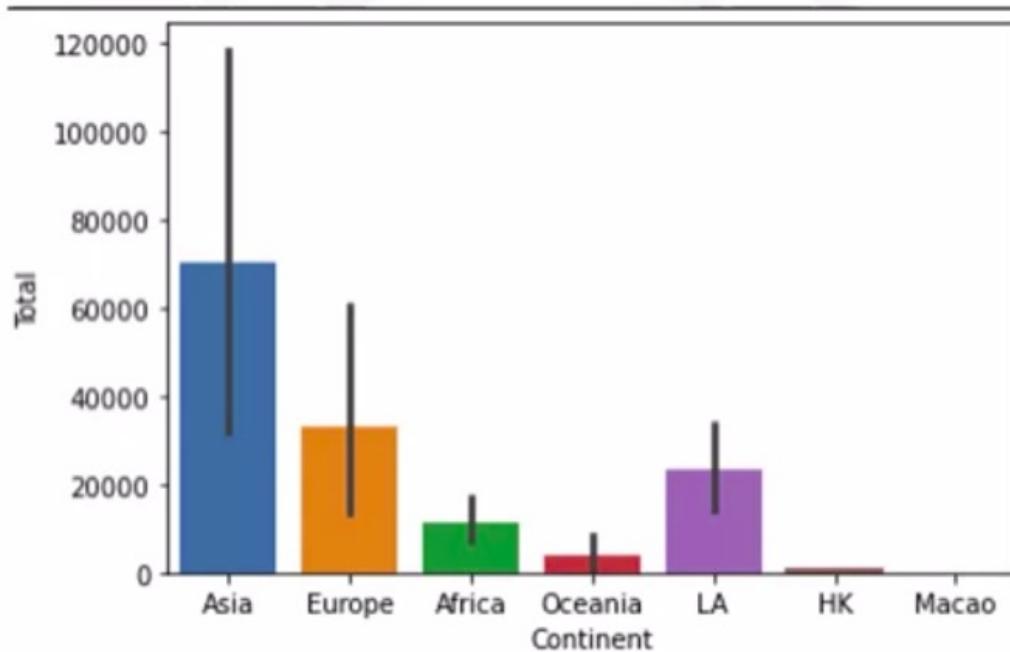
```
sns.countplot(x='Continent', data=df_canada)
```



# Categorical Plots

---

```
sns.barplot(x='Continent',y='Total',data=df_canada)
```



# Recap

---

In this video, you learned that:

- Seaborn is a Data Visualization library based on Matplotlib.
- Seaborn was built primarily to provide a high-level interface for drawing statistical graphics.
- Scatter Plots and Regression Lines can be created with one line of code using Seaborn.
- Seaborn's Regplot function accepts additional parameters for personal customization.

[Back](#)

## Practice Quiz: Advanced Visualization Tools

Practice Quiz • 10 min • 5 total points

## Congratulations! You passed!

Grade received **100%** To pass 80% or higher[Go to next item](#)

1. True or false. Waffle charts are a visualization technique that represents categorical data in the form of square tiles or cells.

1 / 1 point

 True False**Correct**

Correct! Waffle charts are a visualization technique that represents categorical data in the form of square tiles or cells. Their shape has square boxes on a grid that resembles a waffle appearance.

2. What do the squares on a waffle chart represent?

1 / 1 point

- The squares represent the trajectory of the statistics.
- The squares represent location plots on the graph.



Back

## Practice Quiz: Advanced Visualization Tools

Practice Quiz • 10 min • 5 total points

2. What do the squares on a waffle chart represent?

1 / 1 point

- The squares represent the trajectory of the statistics.
- The squares represent location plots on the graph.
- These squares resemble a grid of equal-sized squares, with each square representing a specific value or category.
- These squares resemble a grid of different-sized squares, each representing a specific value or category.

Correct

Correct! The equal-sized squares represent a specific value or category.

3. What is Seaborn based on?

1 / 1 point

- Matplotlib
- Word Clouds
- Python
- Waffle Charts

Correct

Correct! Although Seaborn is another data visualization library, it is based on Matplotlib.

[Back](#) Practice Quiz: Advanced Visualization Tools

Practice Quiz • 10 min • 5 total points

4. Seaborn is a \_\_\_\_\_ visualization library that provides a high-level interface for visualizing geospatial data.

1 / 1 point

- Digital
- Python
- Matplotlib
- Data



Correct

Correct! Seaborn is a Python visualization library based on Matplotlib.

5. The easiest way to create a waffle chart in Python is using the Python package \_\_\_\_\_.

1 / 1 point

- Matplotlib
- PyWaffle
- Waffle chart
- Data visualization library



Correct

Correct! Using the PyWaffle library in Python, you can easily create visually appealing waffle charts to communicate categorical data effectively.

# Introduction to Folium

---

© IBM Corporation. All rights reserved.

# What you will learn

---



Describe Folium and its features



Explain what is Folium used for

# What is Folium?

---

- Folium is a powerful data visualization library in Python that was built primarily to help people visualize geospatial data.
- With Folium, you can create a map of any location in the world using latitude and longitude values. You can also create a map and superimpose markers and clusters on top of the map for interesting visualizations.
- You can also create maps of different styles, such as street-level maps, stamen maps, and a couple of others.

# Creating a world map

```
#import library  
import folium
```

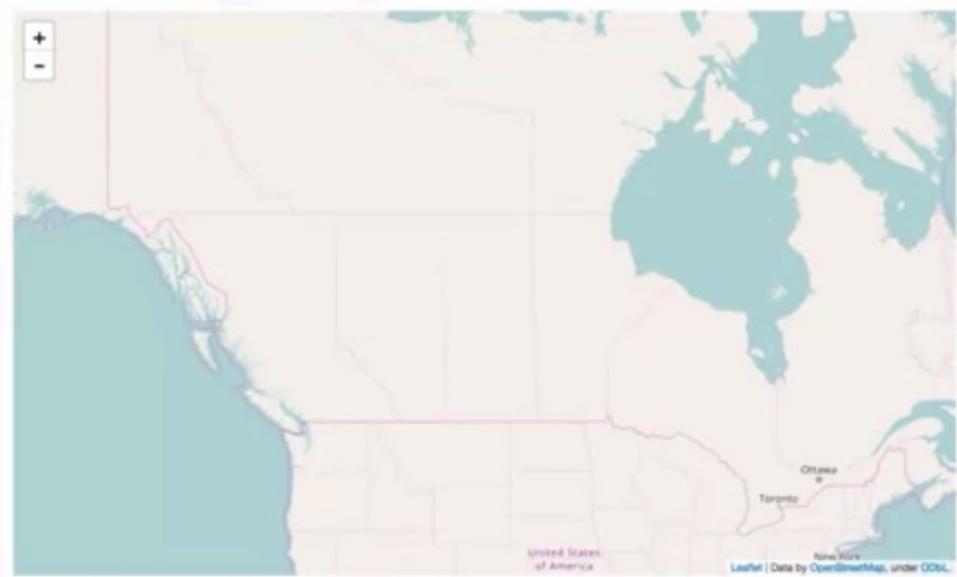
```
# define the world map  
world_map = folium.Map()
```

```
# display world map  
world_map
```



# Creating a map of Canada

```
# define the world map centered around
# Canada with a low zoom level
world_map = folium.Map(
    location=[56.130, -106.35],
    zoom_start=4
)
# display world map
world_map
```



# Map styles: Stamen Toner

```
# create a Stamen Toner map of  
# the world centered around Canada  
world_map = folium.Map(  
    location=[56.130, -106.35],  
    zoom_start=4,  
    tiles='Stamen Toner'  
)  
  
# display map  
world_map
```



# Map styles: Stamen Terrain

```
# create a Stamen Toner map of  
# the world centered around Canada  
world_map = folium.Map(  
    location=[56.130, -106.35],  
    zoom_start=4,  
    tiles='Stamen Terrain'  
)  
  
# display map  
world_map
```



# Recap

---

In this video, you learned that:

- Folium is a data visualization library in Python that helps people visualize geospatial data.
- With Folium, you can create maps of different styles, such as street-level maps, stamen maps, and more.
- A feature of Folium is that you can create different map styles using the tiles parameter.

# Maps with Markers

---

© IBM Corporation. All rights reserved.

# What you will learn

---



Explain how Folium can add markers to a map



Describe how to generate markers

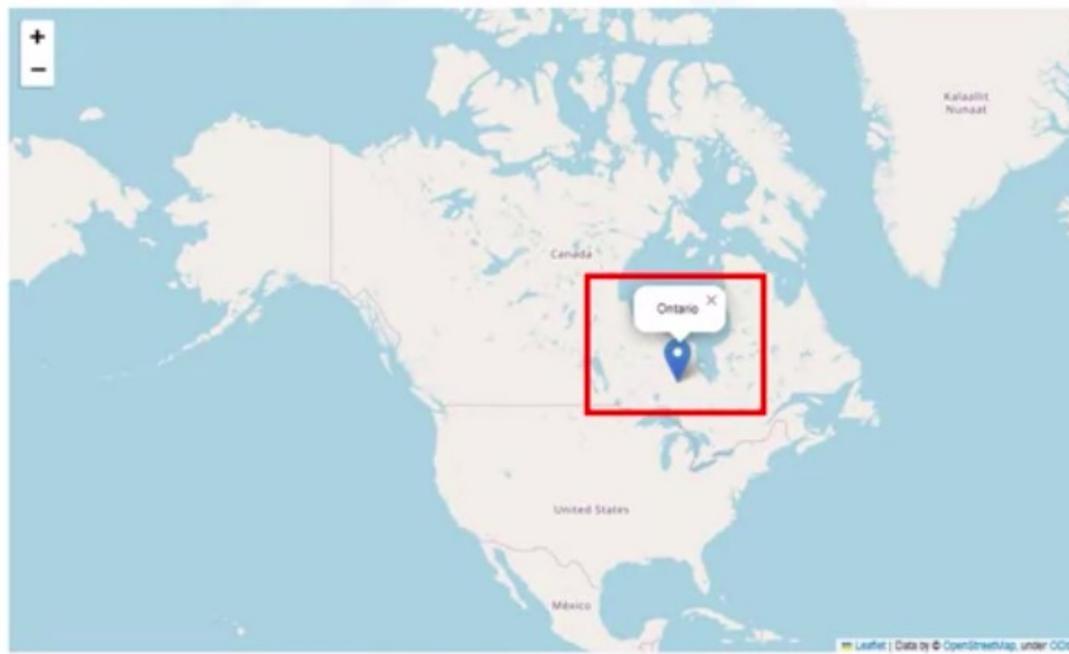
# Initializing map (recap)

```
import folium  
  
# Create a map object centered at Canada  
canada_map = folium.Map(location=[56.1304, -106.3468], zoom_start=4)  
  
# Display the map with the marker  
canada_map
```



# Add marker and label

```
# Add a marker for Ontario province  
folium.Marker(location=[51.2538, -85.3232], popup='Ontario').add_to(canada_map)
```



# Add marker with feature group

```
# generate map of Canada
canada_map = folium.Map(
    location=[56.130, -106.35],
    zoom_start=4
)

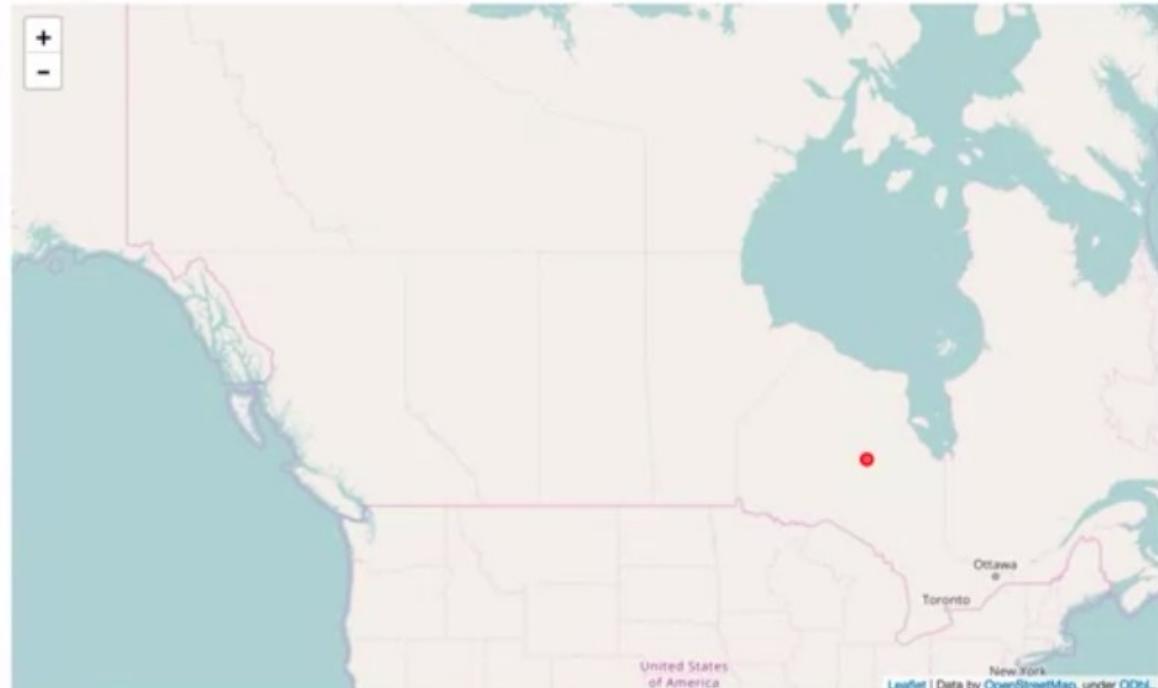
## add a red marker to Ontario

# create a feature group
ontario = folium.map.FeatureGroup()

# style the feature group
ontario.add_child(
    folium.features.CircleMarker(
        [51.25, -85.32], radius = 5,
        color = "red", fill_color = "Red"
    )
)

# add the feature group to the map
canada_map.add_child(ontario)
```

```
# display map
canada_map
```



# Label the marker

```
# generate map of Canada
canada_map = folium.Map(
    location=[56.130, -106.35],
    zoom_start=4
)

## add a red marker to Ontario

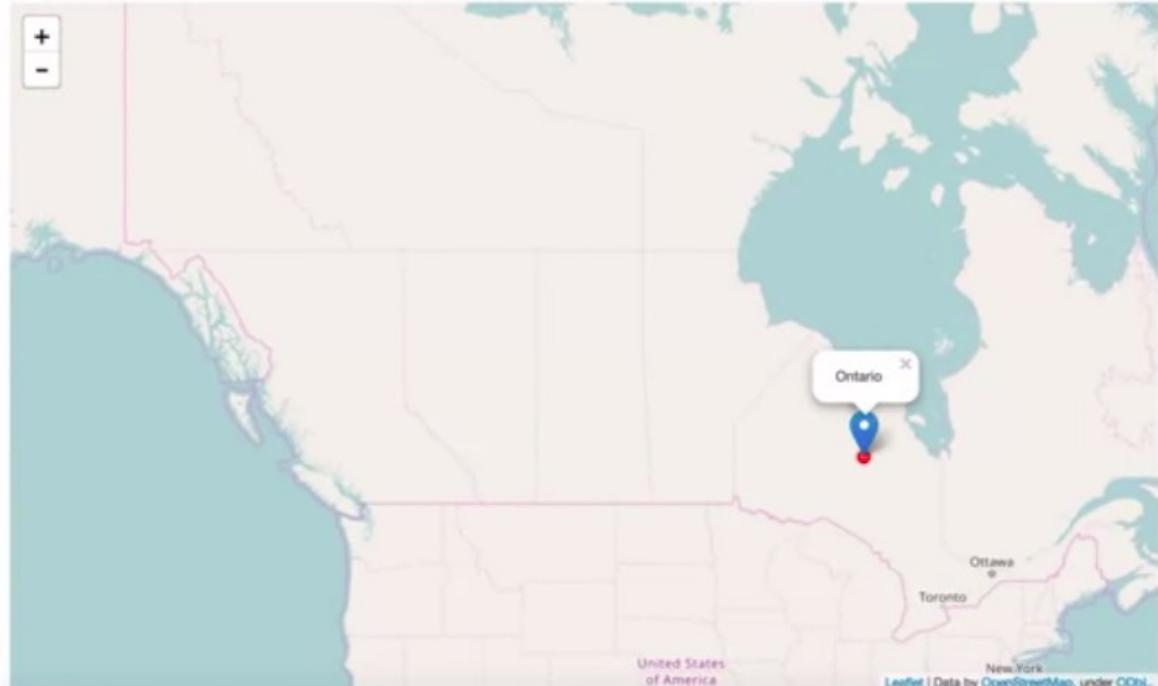
# create a feature group
ontario = folium.map.FeatureGroup()

# style the feature group
ontario.add_child(
    folium.features.CircleMarker(
        [51.25, -85.32], radius = 5,
        color = "red", fill_color = "Red"
    )
)

# add the feature group to the map
canada_map.add_child(ontario)

# label the marker
folium.Marker([51.25, -85.32],
    popup='Ontario').add_to(canada_map)

# display map
canada_map
```

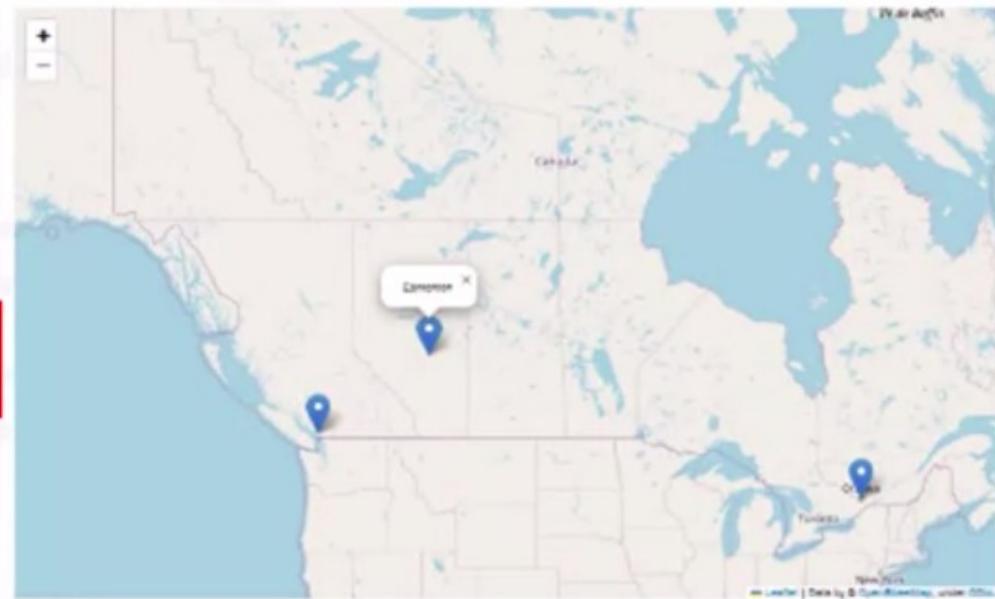


# Multiple markers

```
# Define a list of locations and their corresponding popups
locations = [
    {"location": [45.4215, -75.6989], "popup": "Ottawa"},
    {"location": [53.5461, -113.4938], "popup": "Edmonton"},
    {"location": [49.2827, -123.1207], "popup": "Vancouver"},
    # Add more locations and their popups here
]

# Add markers for each location in the list
for loc in locations:
    folium.Marker(location=loc["location"],
                  popup=loc["popup"]).add_to(map)

# Display the map with the markers
map
```

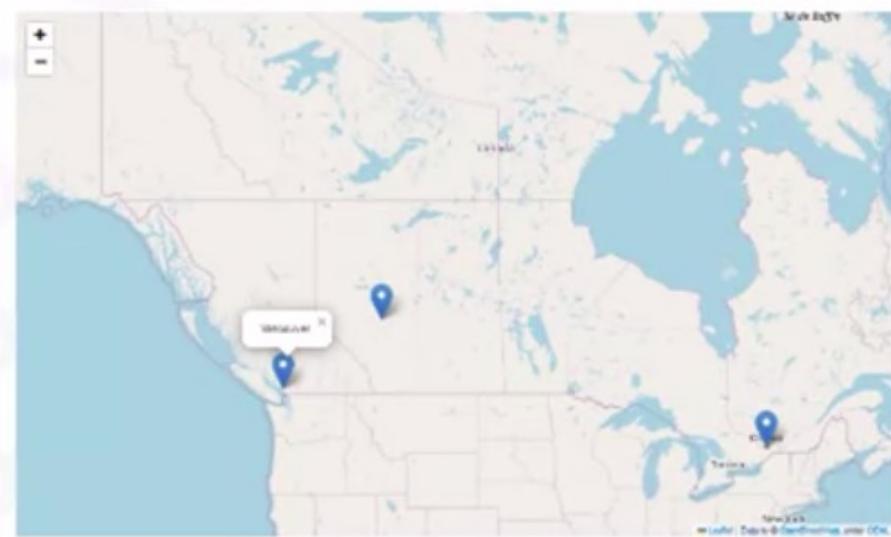


# Multiple markers

```
#import MarkerCluster
from folium.plugins import MarkerCluster

# Create a MarkerCluster object
marker_cluster = MarkerCluster().add_to(map)
```

```
# Add markers for each location in the list to the MarkerCluster
for loc in locations:
    folium.Marker(location=loc["location"],
                  popup=loc["popup"]).add_to(marker_cluster)
```



# Recap

---

In this video, you learned that:

- With Folium, you can easily add markers on maps
- The location parameter specifies the latitude and longitude coordinates of the center point of the map
- Markers play a vital role in enhancing interactivity and adding context to maps
- The folium.Marker() function specifies location parameters
- The popup parameter provides a label upon being clicked
- Markers can be created using “feature group”

# Choropleth Maps

---

© IBM Corporation. All rights reserved.

# What you will learn

---

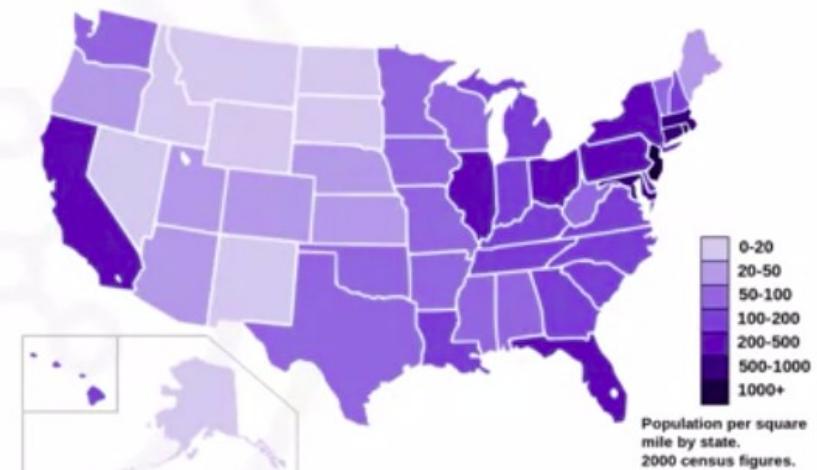
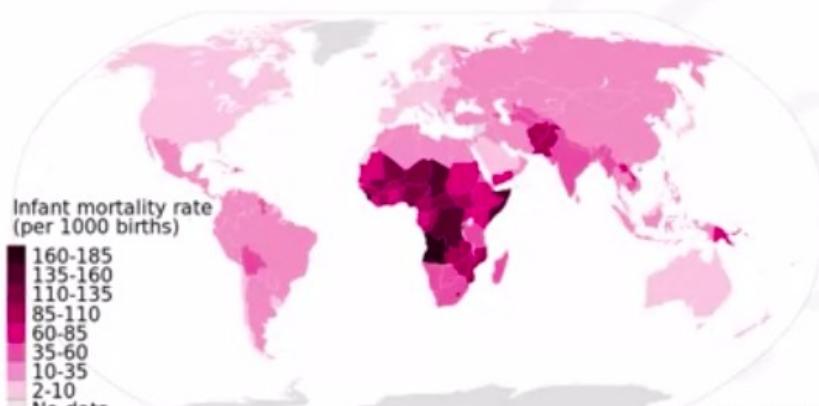


Describe Choropleth Maps



Explain what are the uses for Choropleth Map

# Choropleth Maps



# What is Folium?

---

- Folium is a Python library used for creating interactive maps and visualizations.
- It provides a simple and intuitive way to generate maps using data from various sources, including GeoJSON, Pandas DataFrames, and NumPy arrays.

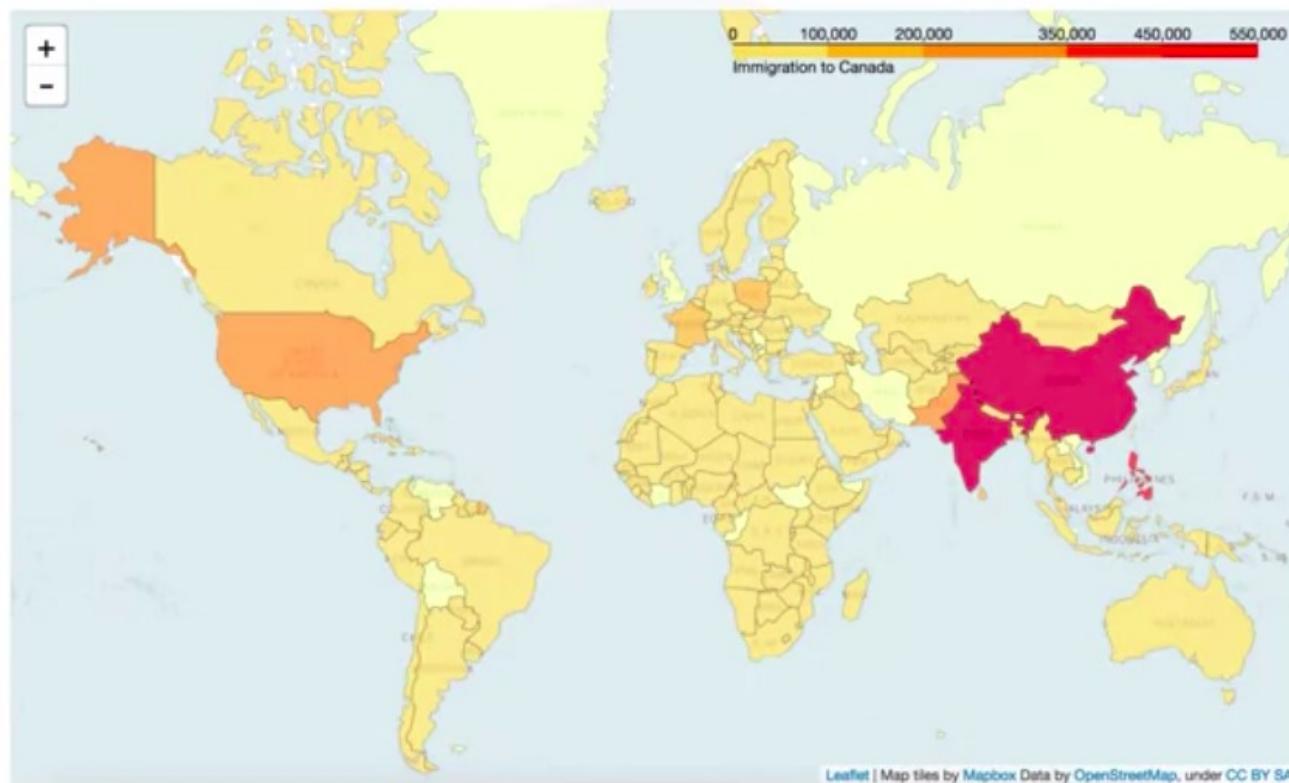
# GeoJson File

---

```
{  
  "type": "FeatureCollection",  
  "features": [  
    {  
      "type": "Feature",  
      "properties": {  
        "name": "Brunei"  
      },  
      "geometry": {  
        "type": "Polygon",  
        "coordinates": [  
          [  
            [114.204017, 4.525874], [114.599961, 4.900011],[115.45071, 5.44773],  
            [115.4057, 4.955228],[115.347461, 4.316636],[114.869557, 4.348314],  
            [114.659596, 4.007637], [114.204017, 4.525874]  
          ]  
        ]  
      },  
      "id": "BRN"  
    },  
  ]}
```

# Creating the Map

---



# Dataset - Processed

	Type	Coverage	Country	AREA	AreaName	REG	RegName	DEV	DevName	1980	...	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
0	Immigrants	Foreigners	Afghanistan	935	Asia	5501	Southern Asia	902	Developing regions	16	...	2978	3436	3009	2652	2111	1746	1758	2203	2635	2004	58639
1	Immigrants	Foreigners	Albania	908	Europe	925	Southern Europe	901	Developed regions	1	...	1450	1223	856	702	560	716	561	539	620	603	15699
2	Immigrants	Foreigners	Algeria	903	Africa	912	Northern Africa	902	Developing regions	80	...	3616	3626	4807	3623	4005	5393	4752	4325	3774	4331	69439
3	Immigrants	Foreigners	American Samoa	909	Oceania	957	Polynesia	902	Developing regions	0	...	0	0	1	0	0	0	0	0	0	6	
4	Immigrants	Foreigners	Andorra	908	Europe	925	Southern Europe	901	Developed regions	0	...	0	0	1	1	0	0	0	1	1	15	

df\_canada

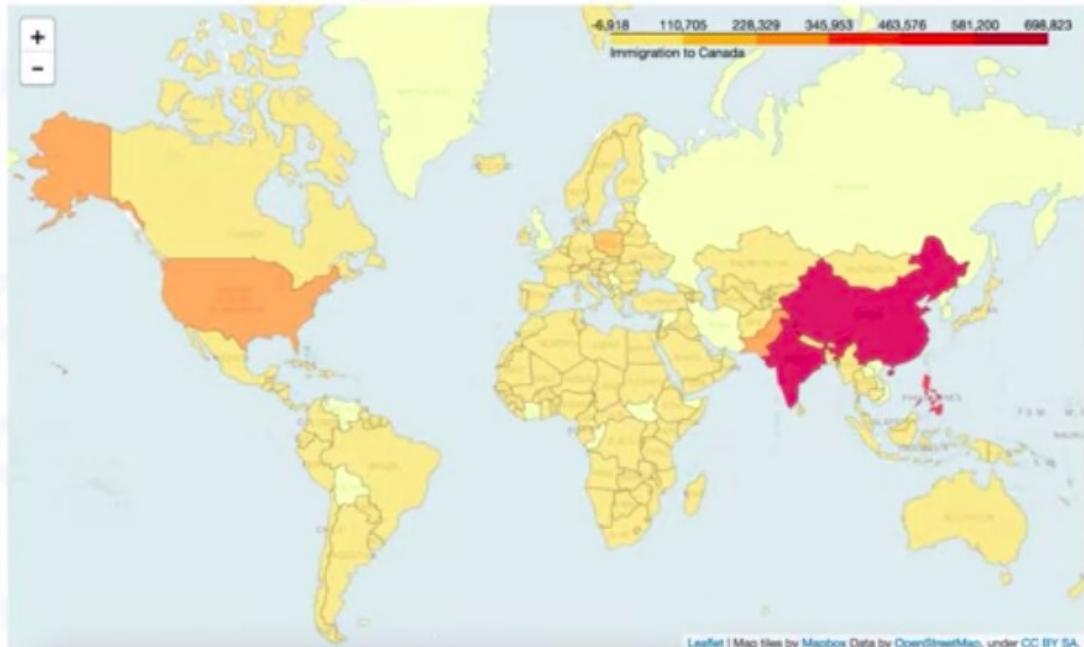
# Creating the Map

```
# create a plain world map
world_map = folium.Map(
    zoom_start=2,
    tiles='Mapbox Bright'
)

## geojson file
world_geo = r'world_countries.json'

# generate choropleth map using the total
# population of each country to Canada from
# 1980 to 2013
world_map.choropleth(
    geo_path=world_geo,
    data=df_canada,
    columns=['Country', 'Total'],
    key_on='feature.properties.name',
    fill_color='YlOrRd',
    legend_name='Immigration to Canada'
)

# display map
world_map
```



# Recap

---

In this video, you learned that:

- A choropleth map is a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable.
- When creating a choropleth map, Folium requires a GeoJson file that includes geospatial data of the region.
- The Mapbox Bright Tileset displays the name of every country when used on a map.

[Back](#) Practice Quiz: Visualizing Geospatial Data  
Practice Quiz • 10 min • 5 total points

## Congratulations! You passed!

Grade received **100%** To pass 80% or higher

[Go to next item](#)

1. Which application is a powerful data visualization library in Python, built primarily to help people visualize geospatial data?

1 / 1 point

- Plotly
- Matplotlib
- PyWaffle
- Folium

 **Correct**

Correct! Folium is an excellent library for geospatial data visualization. It creates interactive and customizable maps.

2. Which Folium map style enables you to visualize natural vegetation colors?

1 / 1 point

- Stamen terrain

[Back](#) Practice Quiz: Visualizing Geospatial Data  
Practice Quiz • 10 min • 5 total points

2. Which Folium map style enables you to visualize natural vegetation colors?

1 / 1 point

- Stamen terrain
- Street-level map
- Open Street map
- Stamen toner



Correct

Correct! The stamen terrain style is great for visualizing hill shading and natural vegetation colors.

3. True or False: With Folium, you can display multiple markers on a map.

1 / 1 point

- True
- False



Correct

Correct! Yes, you can display multiple markers on a map using Folium.

[Back](#) Practice Quiz: Visualizing Geospatial Data

Practice Quiz • 10 min • 5 total points



Correct

Correct! Yes, you can display multiple markers on a map using Folium.

4. Complete the following: To create a choropleth map of a region of interest, Folium requires a Geo JSON file that includes\_\_\_\_\_.

1 / 1 point

- Statistical variable
- Geospatial data of the region
- df\_Canada
- The mapbox bright tiles set



Correct

Correct! For creating a choropleth map of the world, you would need a Geo JSON file that lists each country and any geospatial data to define its borders and boundaries.

5. What is the first step when converting a world map into a choropleth map?

1 / 1 point

- Use the country names to look up the geospatial information
- Use the columns "Country" and "Total" in our df\_Canada dataframe

[Back](#)

## Practice Quiz: Visualizing Geospatial Data

Practice Quiz • 10 min • 5 total points

 df\_Canada The mapbox bright tiles set **Correct**

Correct! For creating a choropleth map of the world, you would need a Geo JSON file that lists each country and any geospatial data to define its borders and boundaries.

5. What is the first step when converting a world map into a choropleth map?

1 / 1 point

 Use the country names to look up the geospatial information Use the columns "Country" and "Total" in our df\_Canada dataframe Define the variable Apply the choropleth function **Correct**

Correct! When converting a world map into a choropleth map, the first step is to define a variable that points to our Geo JSON file.



Search in course

Search



Tushar Raha

Data Visualization with Python > Week 3 > Summary: Advanced Visualizations and Geospatial Data

< Previous Next >

Choropleth Maps  
4 min

Ungraded App Item: Hands-on Lab: Creating Maps and Visualizing Geospatial Data  
30 min

Practice Quiz: Practice Quiz: Visualizing Geospatial Data  
5 questions

Reading: Summary: Advanced Visualizations and Geospatial Data  
5 min

Ungraded Plugin: Cheat Sheet: Maps, Waffles, WordCloud and Seaborn  
15 min

Quiz: Graded Quiz: Advanced Visualizations and Geospatial Data  
10 questions

Congratulations! You have completed this module. At this point in the course, you know:

- Folium is a data visualization library in Python that helps people visualize geospatial data.
- With Folium, you can create maps of different styles, such as street-level maps, stamen maps, and more.
- A feature of Folium is that you can create different map styles using the tiles parameter.
- With Folium, you can easily add markers on maps.
- The 'location' parameter specifies the latitude and longitude coordinates of the center point of the map.
- Markers play a vital role in enhancing interactivity and adding context to maps.
- The folium.Marker() function specifies location parameters.
- The popup parameter provides a label upon being clicked.
- Markers can be created using "feature group."
- A choropleth map is a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable.
- When creating a choropleth map, Folium requires a GeoJson file that includes geospatial data of the region.
- The Mapbox Bright Tileset displays the name of every country when used on a map.



[Back](#)

## Graded Quiz: Advanced Visualizations and Geospatial Data

Due Oct 15, 11:59 PM IST

Graded Quiz • 30 min

## Congratulations! You passed!

Grade received 100% Latest Submission Grade 100% To pass 70% or higher

[Go to next item](#)

1. Although Seaborn is another data visualization library, it is based on \_\_\_\_\_.

1 / 1 point

- Matplotlib
- SciPy
- NumPy
- Pandas



Correct

Correct! Although Seaborn is another data visualization library, it is based on Matplotlib

2. The default map style in Folium is the \_\_\_\_\_.

1 / 1 point

- Arial

[Back](#)

## Graded Quiz: Advanced Visualizations and Geospatial Data

Due Oct 15, 11:59 PM IST

Graded Quiz • 30 min

Correct

Correct! Although Seaborn is another data visualization library, it is based on Matplotlib

2. The default map style in Folium is the \_\_\_\_\_.

1 / 1 point

- Arial
- Stamen Toner
- Open Street Map
- Stamen Terrain

Correct

Correct! The default map style in Folium is the Open Street Map. It shows a street view of an area when you are zoomed in and the borders of the world countries when you are zoomed out all the way.

3. What is the code for setting the initial zoom level in Folium?

1 / 1 point

- zoom-start parameter
- zoom\_start=parameter

[Back](#)

## Graded Quiz: Advanced Visualizations and Geospatial Data

Graded Quiz • 30 min

Due Oct 15, 11:59 PM IST

3. What is the code for setting the initial zoom level in Folium?

1 / 1 point

- zoom-start parameter
- zoom\_start=parameter
- {zoom\_start} parameter
- (zoom\_start) parameter

 Correct

Correct! With the initial zoom, you can easily change the zoom level after the map is rendered by zooming in or out. You can play with this parameter to determine the initial zoom level for different values.

4. True or False. The 'location' parameter specifies the latitude and longitude coordinates of the map's center point.

1 / 1 point

- True
- False

 Correct

Correct! The 'location' parameter specifies the latitude and longitude coordinates of the map's center point.

[Back](#) Graded Quiz: Advanced Visualizations and Geospatial Data

Due Oct 15, 11:59 PM IST

Graded Quiz • 30 min

5. \_\_\_\_\_ represent specific locations or points of interest, providing additional information when clicked.

1 / 1 point

- Locations
- Markers
- Navigation
- Zoom function



Correct

Correct! Markers represent specific locations or points of interest, providing additional information when clicked. Markers are like signposts that guide us through the map, highlighting essential elements.

6. True or False. A choropleth map is a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable displayed on the map.

1 / 1 point

- False
- True



Correct

Correct! A choropleth map is a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable displayed on the map.

[Back](#) Graded Quiz: Advanced Visualizations and Geospatial Data

Due Oct 15, 11:59 PM IST

Graded Quiz • 30 min

7. True or False. Folium is a Python library used for creating interactive maps and visualizations.

1 / 1 point

- True  
 False

 Correct

Correct! Folium is a Python library used for creating interactive maps and visualizations. It provides a simple and intuitive way to generate maps using data from various sources, including GeoJSON, Pandas DataFrames, and NumPy arrays.

8. What kind of file does Folium require to create a Choropleth map of a specific region?

1 / 1 point

- Json  
 GeoJson  
 Geo  
 HTML

 Correct

Correct! To create a choropleth map of a region of interest, Folium requires a GeoJson file that includes geospatial data of the region.

[Back](#)

## Graded Quiz: Advanced Visualizations and Geospatial Data

Due Oct 15, 11:59 PM IST

Graded Quiz • 30 min

9.

1 / 1 point

Which of the following are specialized plot types in Seaborn? *Select two.*

 Categorical plot

Correct

Correct! Seaborn provides specialized plot types such as regression, distribution, and categorical plots that are particularly useful for analyzing data and modeling relationships.

 Bar plot Pie plot Regression plot

Correct

Correct! Seaborn provides specialized plot types such as regression, distribution, and categorical plots that are particularly useful for analyzing data and modeling relationships.

10. What type of data do Waffle Charts represent?

1 / 1 point

[Back](#)

## Graded Quiz: Advanced Visualizations and Geospatial Data

Due Oct 15, 11:59 PM IST

Graded Quiz • 30 min

 Pie plot Regression plot **Correct**

Correct! Seaborn provides specialized plot types such as regression, distribution, and categorical plots that are particularly useful for analyzing data and modeling relationships.

**10.** What type of data do Waffle Charts represent?

1 / 1 point

- Sequential
- Categorical
- Numerical
- Statistical

**Correct**

Correct! Waffle charts are a visualization technique that represents categorical data in the form of square tiles or cells.